

Chapter One: Contents

(Generalized Mode Choice – 10 July 2002 – LA-UR-02-4217– TRANSIMS 3.0)

1. INTRODUCTION	2
2. GENERAL METHODOLOGY	4
2.1 MODE CHOICE THEORY	4
2.2 MODE CHOICE USING LOGITS	13
2.3 A SECOND MODE CHOICE METHOD	16
3. SUMMARY	25

Chapter Two: Figures

<i>Fig. 1. A transit trip where the origin and destination are both within ¼ mile of a transit stop may take longer than walking between the two activities. The TRANSIMS Route Planner, in these cases, will have the traveler walk rather than take transit.</i>	<i>2</i>
<i>Fig. 2. The relationship between the transit and auto “costs” for two different sets of calibration parameters (α). Values of α are determined so that the proportion of “costs” below the line match the mode split for transit. The arrow shows the general direction the points moved as the values of α changed.</i>	<i>19</i>
<i>Fig. 3. Two techniques for stochastic mode selection. Panel (a) shows the method when both “costs” are computed. Estimates of mode selection using panel (b) require only generalized transit “costs.”</i>	<i>20</i>
<i>Fig. 4. Auto costs for transit and walk calibrations plotted against each other. The black points are the calibration data where the Route Planner produces all walk legs. The red points represent those where the route from the Route Planner has at least one transit leg. In an application, the auto costs $C = (c_a^t, c_a^w)$ are computed, and $P(s_w C)$ is estimated by the proportion of black dots in the circle.</i>	<i>23</i>

Chapter Two: Tables

Table 1. The possible outcomes from the Route Planner for specified modes.	3
Table 2. The mode symbols used in this section.	5
Table 3. Conditions when $P_R(m M, I) = 0$.	6
Table 4. Conditions when $P_R(m M, I) = 1$.	6

Chapter One—Generalized Mode Choice

1. INTRODUCTION

In TRANSIMS, the Route Planner, when given a mode string, will find a route that is the fastest possible. Travel by each “mode” in TRANSIMS is proceeded by and ends with a walk. Travel by transit in TRANSIMS is denoted by the Route Planner as wtw, where w stands for walk and t stands for transit. The walks allow the traveler to move from a parking location, an activity location, or another transit stop to the desired transit stop. At first glance, this should have no impact on a mode-calibrated activity set, but it does. Suppose the mode string wtw (transit) is requested to move a traveler between two activity sites. The traveler will walk from the first activity to find a transit route to move him to a transit stop close to the second activity and walk from this transit stop to the second activity. If the Route Planner, in its search, finds a walking route that is faster than the combined two walks and a transit ride, the route will be completed entirely by walking. Such a trip is diagramed in Fig. 1 where the walks to and from the transit stop are $\frac{1}{4}$ mile each. The time to make these walks, wait at the stop for the transit vehicle, and ride transit is more than it takes to walk between the two activity sites.

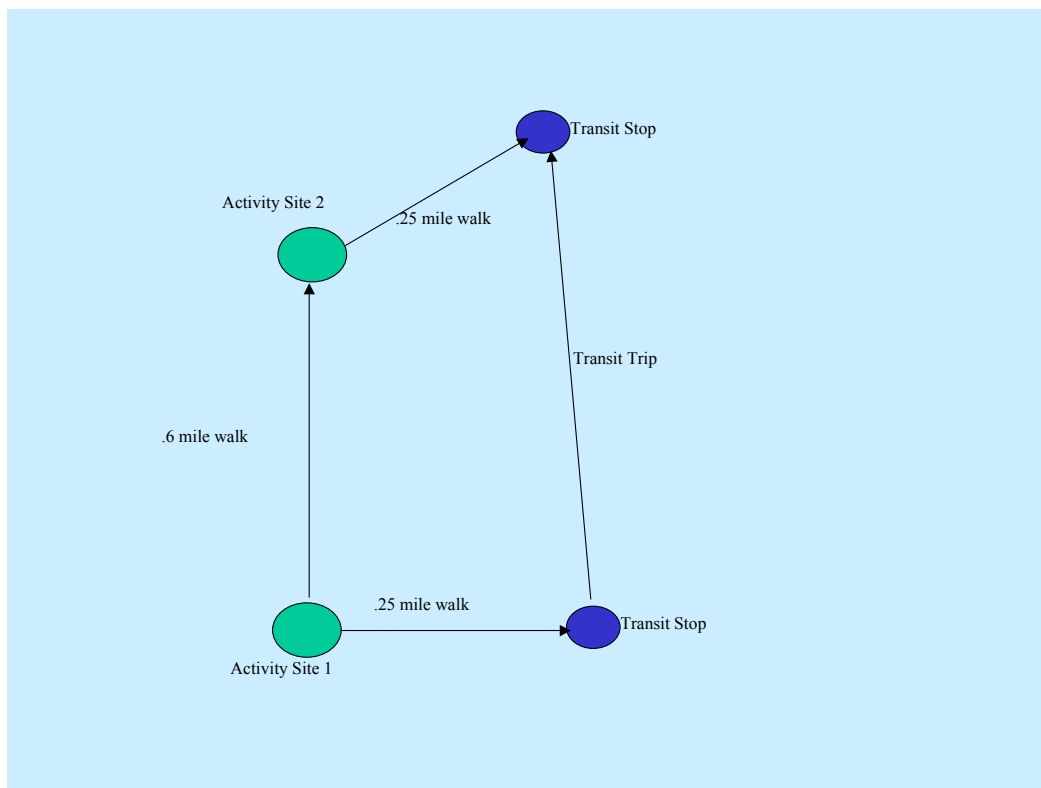


Fig. 1. A transit trip where the origin and destination are both within $\frac{1}{4}$ mile of a transit stop may take longer than walking between the two activities. The TRANSIMS Route Planner, in these cases, will have the traveler walk rather than take transit.

Table 1 shows the possible outcomes from the Route Planner for some of the modes. For each mode, one of the outcomes is “not possible”. This occurs when walks become longer than the limits set in the Route Planner configuration file keys, or there is no walking path between two activity locations. For drive trips, this outcome indicates that on previous activities the traveler moved away from his vehicle, and there is no walking path back to it.

Table 1. The possible outcomes from the Route Planner for specified modes.

Mode	Result
walk	walk, “not possible”
drive	walk, walk-car-walk, “not possible”
transit	walk, any combination of walk..bus..light rail..walk, “not possible
light rail	walk, walk-light rail-walk, “not possible”

One should note in Table 1 that a possible outcome for each of the modes is *walk*. This is true even with drive trips and will happen if the origin and destination of the trip are across the street from one another. To the Route Planner, transit trips are any combinations of walks, buses, and light rail. Light rail trips are any combination of light rail and walk. In both of these cases, a bus or light rail leg on the trip is not necessary.

It is apparent that a mode choice methodology using the TRANSIMS technology must be calibrated and implemented differently than is usually done in other transportation modeling systems. A perfectly calibrated logit function for mode choice among, say, bus, walk, and light rail, when applied to an activity set, will no longer be in calibration after the Route Planner routes the travelers. This is caused by the aggregate nature of the logit function. A logit is a statistical fit to external data that “smooths” over multiple individuals with the same demographic and travel characteristics. This smoothing does not fit each individual, but estimates the tendency of individuals of that type to drive—take transit, etc. In TRANSIMS, each individual never loses its identity. Results of the simulation are aggregated after the activities have been set, the trips routed, and the entire set of trips individually (in the presence of all other travelers) simulated.

From the above discussion, the usual mode choice logit methodology cannot be applied directly to an activity set and maintain its calibration. However, mode choice logits derived from the survey data or other sources may be used in calibrating a model built with the TRANSIMS technology. Non-logit methods may also be used.

The purpose of this document is to provide the mathematical justification for TRANSIMS mode methodologies. TRANSIMS-compatible logit methodologies are given along with a non-logit methodology based on general cost functions.

2. GENERAL METHODOLOGY

Mode choice in TRANSIMS is accomplished using external functions, such as logits and travel cost functions, along with a meta-method using the TRANSIMS Framework. The exact mode choice methodology in any TRANSIMS study is left to the analyst. This section outlines some methodologies for mode choice that may be implemented using the TRANSIMS technology.

It is clear from the previous section that care must be taken when calibrating the modes in TRANSIMS. The Route Planner chooses routes, and hence modes, based on best network travel times rather than the explicit mode requirement from the activity list. This section gives two methods for calibrating modes given this behavior of the Route Planner. Many other methodologies could be developed for this task. However, Route Planner characteristics must be taken into account when these methodologies are developed.

Mathematics for two mode choice methodologies is given in this section. The first methodology is based on logit fits to survey and other data. These logit fits must take into account the “best time” behavior of the Route Planner. The second methodology uses the survey data to fix the proportional mode splits and some user-defined, predetermined cost functions. Here, the model is calibrated by matching the mode splits in the synthetic population to the target mode splits.

All mode choice procedures have two steps. First, a calibration method is used to guarantee the proper mode splits in the calibration or base year. Then, this calibration function is applied to the forecast year.

A third step may be required in TRANSIMS or any other truly disaggregated transportation modeling system. In these systems, synthetic individuals are not only given a mode to reach an activity in the activity list, but also are given a link-by-link and transit route-by-transit route plan to move them from one activity location to another. In many cases, these travel plans may be considered unrealistic, such as trips with multiple transfers between transit lines and unreasonable walks between the transit stops. Because of the aggregate nature of any statistically based methodology, including logits, human “behavior” is duplicated on a population rather than an individual basis. Therefore, some individuals in the synthetic population will be assigned modes that seem unrealistic. In TRANSIMS, these are corrected by iteration between the Activity Regenerator and the Route Planner.

2.1 Mode Choice Theory

TRANSIMS, being a completely disaggregated system, requires care in the calibration and application of mode choice. This and the subsequent subsections give the general mathematical and application theory of mode choice in TRANSIMS. It is critical for correct mode assignment that this theory is understood. Sections 2.1.1, 2.1.2 and 2.1.3 outline the general mathematics of mode choice and give calibration and application mode choice theory.

Uppercase letters for modes are used in this section to indicate the mode in the activity list. Lowercase letters denote Route Planner modes. These are used in the Traffic Microsimulator.

2.1.1 Preliminary Mathematics

Let a set of demographic and “land-use” characteristics, such as age, income, and distance to the nearest transit stop, for each individual in the population be denoted by $S = \{I_1, I_2, \dots, I_m\}$. The probability, $P_T(m|I)$, that mode m is used in the Traffic Microsimulator for a traveler/trip with characteristic I is given by

$$P_T(m|I) = \sum_M P_R(m|M, I) P_U(M|I) \quad (1)$$

where $P_R(m|M, I)$ is the probability that the Route Planner assigns mode m given that the activity list assigns mode M . This probability is a function of the “best travel time” found by the Route Planner. $P_U(M|I)$ is a user-defined function that assigns mode M in the activity file.

Throughout this section, the symbols in Table 2 are used. Here, lowercase letters indicate Route Planner modes, while uppercase letters are the modes assigned by the Activity Generator.

Table 2. The mode symbols used in this section.

Route Symbol (m)	Activity Symbol (M)	Router Modes Given the Activity Mode
w	W	Walk
a	A	Auto, Possible walk
t	T	Any combination of walk, bus, light rail
b	B	Any combination of walk, bus
l	L	Any combination of walk, light rail
bi	BI	Bike
m	M	Inter-household shared ride
s	S	School bus
NA	X	Any of the symbols w, a, t, b, l

From Table 2 it is apparent that the probability $P_R(m|M, I)$ has fixed values for some combinations of m and M . Some pairs of m and M never occur. Table 3 shows conditions when $P_R(m|M, I) = 0$.

Table 3. Conditions when $P_R(m|M, I) = 0$.

m=	M=
a	Not A
bi	Not BI
m	Not M
s	Not S

Also, pairs of m and M that are guaranteed happen with probability 1. Table 4 shows conditions when $P_R(m|M, I) = 1$.

Table 4. Conditions when $P_R(m|M, I) = 1$.

m=	M=
bi	BI
s	S
m	M
w	W

Using Tables 3 and 4 and Equation 1, the following relationships are derived for each of the TRANSIMS modes. These equations assume that there is only a small probability that the Route Planner will find a walking path that is faster than driving when the activity calls for a drive trip. This occurs in the rare cases when the Activity Generator, because of its aggregate nature, chooses activity locations across the street from one another. Equations relating the TRANSIMS final mode probabilities to the Route Planner “rational choice” mode probabilities and a user-defined mode choice probability used in the activity set follow.

For bikes:

$$\begin{aligned}
 P_T(m = bi|I) &= \sum_M P_R(m = bi|M, I)P_U(M|I) \\
 &= P_R(m = bi|M = BI, I)P_U(M = BI|I) \quad (2) \\
 &= P_U(M = BI|I)
 \end{aligned}$$

For school buses:

$$\begin{aligned}
 P_T(m = s|I) &= \sum_M P_R(m = s|M, I)P_U(M|I) \\
 &= P_R(m = s|M = S, I)P_U(M = S|I) \quad (3) \\
 &= P_U(M = S|I)
 \end{aligned}$$

For inter-household shared rides:

$$\begin{aligned}
 P_T(m = m|I) &= \sum_M P_R(m = m|M, I)P_U(M|I) \\
 &= P_R(m = m|M = M, I)P_U(M = M|I) \quad (4) \\
 &= P_U(M = M|I)
 \end{aligned}$$

For autos:

$$\begin{aligned}
 P_T(m = a|I) &= \sum_M P_R(m = a| M, I)P_U(M|I) \\
 &= (1 - P_R(m = w| M = A, I))P_U(M = A|I) \quad (5) \\
 &\approx P_U(M = A|I)
 \end{aligned}$$

For walks:

$$\begin{aligned}
 P_T(m = w|I) &= \sum_M P_R(m = w| M, I)P_U(M|I) \\
 &= P_R(m = w| M = W, I)P_U(M = W|I) + P_R(m = w| M = B, I)P_U(M = B|I) + \quad (6) \\
 &P_R(m = w| M = L, I)P_U(M = L|I) + P_R(m = w| M = T, I)P_U(M = T|I) + \\
 &P_R(m = w| M = A, I)P_U(M = A|I)
 \end{aligned}$$

For buses:

$$\begin{aligned}
 P_T(m = b|I) &= \sum_M P_R(m = b| M, I)P_U(M|I) \\
 &= P_R(m = b| M = T, I)P_U(M = T|I) + (1 - P_R(m = w| M = B, I))P_U(M = B|I) \quad (7)
 \end{aligned}$$

For light rail:

$$\begin{aligned}
 P_T(m = l|I) &= \sum_M P_R(m = l| M, I)P_U(M|I) \\
 &= P_R(m = l| M = T, I)P_U(M = T|I) + (1 - P_R(m = w| M = L, I))P_U(M = L|I) \quad (8)
 \end{aligned}$$

Equations 6, 7, and 8 become simpler if the activity set modes W , B , and L are all replaced with the activity set mode T . That is, $P_U(M = W|I)$, $P_U(M = B|I)$, and $P_U(M = L|I)$ are set to 0 in the user-defined choice function,. In this case, Equations 6, 7, and 8 become:

For walks:

$$\begin{aligned}
 P_T(m = w|I) &= \sum_M P_R(m = w| M, I)P_U(M|I) \\
 &= P_R(m = w| M = T, I)P_U(M = T|I) + P_R(m = w| M = A, I)P_U(M = A|I) \quad (9) \\
 &\approx P_R(m = w| M = T, I)P_U(M = T|I)
 \end{aligned}$$

For buses:

$$\begin{aligned}
 P_T(m = b|I) &= \sum_M P_R(m = b| M, I)P_U(M|I) \\
 &= P_R(m = b| M = T, I)P_U(M = T|I) \quad (10)
 \end{aligned}$$

For light rail:

$$\begin{aligned} P_T(m = l|I) &= \sum_M P_R(m = l|M, I)P_U(M|I) \\ &= P_R(m = l|M = T, I)P_U(M = T|I) \quad (11) \end{aligned}$$

Comparing Equations 9, 10, and 11 with Equations 6, 7, and 8, it is clear that using T , the general transit mode that includes walks, greatly reduces the complexity of transit (including walks) calibrations. When the mode T is used in the activity set, the Route Planner makes a “rational” choice based on travel times between the walk, bus, and light rail modes.

The probability $P_U(M|I)$ is a user-defined probability function that assigns modes in the Activity Generator for individuals/trips with characteristics denoted by I . In non-TRANSIMS applications, $P_U(M|I)$ is usually a calibrated logit function. It is apparent from Equations 6, 7, and 8 that the modes walk, bus, and light rail assigned in the Activity Generator using a calibrated logit function may no longer be in calibration after routing in TRANSIMS. In most cases, the Route Planner in TRANSIMS replaces short bus or light rail trips with walks. Short walks rather than short bus or light rail trips are a realistic representation of personal travel behavior that is difficult to capture using a standard logit approach because logits are more aggregate in nature. This characteristic of TRANSIMS requires that different methods be developed for mode calibration. Mode calibration in TRANSIMS calls for a determination of $P_U(M|I)$ to calibrate the mode choice of the activity set. It also requires that the behavior of the Route Planner, $P_R(m|M, I)$, be accounted for in the calibrations.

The next subsection gives some mathematical considerations for developing mode calibration methodologies and procedures in TRANSIMS. These procedures take into account both the assignment of the mode in the activity set, $P_U(M|I)$, and the “rational choice” behavior of the Route Planner for some of the modes, $P_R(m|M, I)$.

The mathematics given in this document show methods for calibration of the Route Planner “rational choice” between walk, light rail, and bus. This partition of the mode calibration is for demonstration purposes only. The modes may be calibrated using any mutually exclusive partition of the mode choice induced by the Route Planner. For example, the trips/tours could be partitioned into two groups—those that have only walk legs and those that have at least one transit leg. Other partitions could be those that have all walk legs, some bus but no light rail legs, some light rail but no bus legs, and those that have mixed light rail and bus legs. One could also consider the number of transit transfers when deciding on the partitions. Of course, whatever the partitioning scheme is, data must be available to calibrate each of the partitions.

2.1.2 Calibrating Mode Choice

All mode calibration methodologies are aggregated procedures. Mode splits and their associated estimated parametric or nonparametric forms are obtained across many individuals to maintain the base year mode split proportions. Therefore, the estimation of

calibration parameters or functions is done on groups of individuals rather than an individual basis. These groups could represent the entire population or a subgroup of the population, such as those persons with trips that begin and end close to transit stops. In this section, the calibration group is denoted by D , and D replaces the characteristics of the individual (I) in the previous equations.

The term $P_U(M|I)$ in Equation 1 is a user-defined function used for mode calibration.

The structure of this function and other user-defined functions is the basis for mode calibration in TRANSIMS. These functions are investigated in this subsection. Here, and throughout the remainder of this calibration subsection, D replaces I in the functions to denote a group of individuals or a randomly chosen individual from a group with characteristics D .

Many theoretical and mathematically correct general mode methodologies could be developed based on Equations 1 through 11. Only one is given here. The Activity Generator modes, W , L , and B , are replaced with the general transit mode (T). The calibration is carried out in two stages. For the first stage of the calibration, modes A and T are combined into a single “mode” denoted by X . Let $P_U^{(1)}(M_1|D)$ be a calibration function for choosing mode M_1 equal to BI , M , S , or X . See Table 2 for a definition of these modes. It should be noted that M_1 could take on value for modes not listed here as long as the Route Planner does not change it. Then

$$P_U(M|D) = P_U^{(1)}(M_1|D) \text{ for } M_1 = BI, M, S$$

$$P_U(M = T|D) = P_U^{(2)}(M_2 = T|M_1 = X, D)P_U^{(1)}(M_1 = X|D) \quad (12)$$

and

$$P_U(M = A|D) = P_U^{(2)}(M_2 = A|M_1 = X, D)P_U^{(1)}(M_1 = X|D)$$

In Equation 12, $P_U^{(2)}(M_2|M_1 = X, D)$ is a user-defined probability function for choosing mode $M_2 = A$ or T given that mode X is picked in the first step. This second stage calibration, $P_U^{(2)}(M_2|M_1 = X, D)$, must take into account that the Route Planner will make a “rational choice” among walk, bus, and light rail given the activity mode is T .

As a shorthand for what follows, let $P_U^{(2)}(M_2|D)$ represent $P_U^{(2)}(M_2|M_1 = X, D)$ where $M_2 = A$ or T , and it is understood that the probability is conditional on $M_1 = X$. Also, let s_w, s_b, s_l represent the mutually exclusive events that the Route Planner chooses—walk, bus, or light rail, respectively—for a randomly selected trip conditional on the event that the general transit mode (T) is requested. Here, the events s_w, s_b, s_l represent trips with walk only legs, trips with at least one bus leg, and trips with at least one light rail but no bus legs, respectively. However, this is for demonstration purposes only. The only requirement on the events s_w, s_b, s_l is that they are mutually exclusive, and other mutually exclusive partitions of these events are acceptable.

Using Equations 9 through 12 and conditional on the first choice being $M_1 = X$ and that $M_2 = A$ or T :

$$\begin{aligned} P_U^{(2)}(M_2 = A|D) &= P_U^w(M_2 = A|s_w, D)P_R(m = w|M = T, D) + \\ &P_U^b(M_2 = A|s_b, D)P_R(m = b|M = T, D) + \quad (13) \\ &P_U^l(M_2 = A|s_l, D)P_R(m = l|M = T, D) \end{aligned}$$

The probability functions $P_U^k(M_2 = A|s_k, D)$ for $k=w, l, b$ are user-defined calibration functions for assigning the auto mode (A) in the activity list. These probabilities assume that the first choice is $M_1 = X$, and the Route Planner classifies the trip as belonging to s_k , $k=w, l, b$, when given the general transit mode (T).

In the following, let $p_R^k = P_R^k(m = k|M = T, D)$ for $k=w, l, b$. Note that $p_R^w + p_R^b + p_R^l = 1$.

For calibration, let a_c , w_c , b_c , and l_c be calibration target numbers from a base year reflecting the number of auto, walk, bus, and light rail trips for the category denoted by D . Then, a method to attain calibration between these modes is to let

$$P_U^w(M_2 = A|s_w, D) = 1 - \frac{w_c}{(a_c + w_c + b_c + l_c)p_R^w} \quad (14)$$

$$P_U^b(M_2 = A|s_b, D) = 1 - \frac{b_c}{(a_c + w_c + b_c + l_c)p_R^b} \quad (15)$$

and

$$P_U^l(M_2 = A|s_l, D) = 1 - \frac{l_c}{(a_c + w_c + b_c + l_c)p_R^l} \quad (16)$$

Equations 14 to 16 can be set if, and only if, the following relationships hold:

$$\begin{aligned} \frac{w_c}{(a_c + w_c + b_c + l_c)} &\leq p_R^w \\ \frac{b_c}{(a_c + w_c + b_c + l_c)} &\leq p_R^b \quad (17) \\ \frac{l_c}{(a_c + w_c + b_c + l_c)} &\leq p_R^l \end{aligned}$$

It would be unusual for the relationships in Equation 17 not to hold. If, for example,

$\frac{w_c}{(a_c + w_c + b_c + l_c)} > p_R^w$, then, the proportion of walk trips relative to the drives, buses, and light rails in the calibration data is greater than the proportion of walks relative to the buses and light rail (auto not included) from the Route Planner. This would indicate completely irrational behavior on a large segment of the population, an unlikely event.

Inserting Equations 14 to 16 into Equation 13 shows that, in the context of Equation 1:

For drive trips:

$$P_T(m = a | M_1 = X, D) = \frac{a_c}{a_c + w_c + b_c + l_c}$$

For walks:

$$\begin{aligned} P_T(m = w | M_1 = X, D) &= P(s_w \text{ and } M_2 \neq A | M_1 = X, D) \\ &= P(s_w)P(M_2 \neq A | s_w, M_1 = X, D) \\ &= P_R(m = w | M_2 = T, M_1 = X, D)(1 - P_U^{(2)}(M_2 = A | s_w, M_1 = X, D)) \\ &= P_R^w \frac{w_c}{(a_c + w_c + b_c + l_c)p_R^w} = \frac{w_c}{(a_c + w_c + b_c + l_c)} \end{aligned}$$

Similar equations hold for buses and light rail.

The calibration equations given in this subsection represent but a single way to develop mode choice methodologies in TRANSIMS. General methods could be developed using Equations 6, 7, and 8 rather than Equations 9, 10, and 11. The first stage calibration function, $P_U^{(1)}(M_1 | D)$, could include both modes A and T , rather than the combined “mode” X . Such a calibration applied to a forecast year would rely on the “rational choice” behavior of the Route Planner when given mode T . That is, given an assignment of mode T , the traveler will pick walk, light rail, bus, or a combination of these that takes minimum time. Since, the choice of T is calibrated to the proportion of travelers walking and using buses or light rail, the choice of the Route Planner in splitting these modes is probably more close to the truth than an aggregated calibration fit.

2.1.3 Applying Mode Choice to a Forecast Year

This subsection describes the application of calibrated mode choice functions in a forecast setting. It is assumed that the calibration procedure follows that in Section 2.2 and that the functions $P_U^{(1)}(M_1 | D)$, $P_U^w(M_2 = A | s_w, D)$, $P_U^b(M_2 = A | s_b, D)$, and $P_U^l(M_2 = A | s_l, D)$ are calibrated to the base year. The modes, M_l , include the combined auto, walk, bus, and light rail modes (X), BI , M , and S .

All methods of forecasting, whether they be modes, economics, etc., assume that the forecast year behavior, when given the characteristics of the surrounding infrastructure, is the same as that captured in the base year calibrations. In forecasting future year transportation systems, the roadway, the level of transit service, and the population are among the many things bound to be different from those in the base year. These changes are assumed to be reflected in the base year calibration function in the form of changed travel times, transit fares, parking costs, etc. A personal choice when given these new “costs” may change from the choice with the base year “costs”, but personal behavior remains the same when the “costs” do not change. For example, a person making a drive trip in the base year may switch to a transit trip in the future year because of an increased

level of service, increased parking costs, or reduced fares. However, if there are no changes in the underlying “costs,” driving remains the rational choice for the traveler.

Any user-defined mode calibration functions for TRANSIMS must make the same basic assumptions about behaviors. As stated above, this is the standard assumption concerning forecast methodologies. In addition, changes in the infrastructure and populations in the forecast year will cause a different split in the probabilities p_R^w, p_R^b, p_R^l . For example, if the level of transit service increases, the proportion of walks would most likely decrease. This is expected and represents a “rational choice” made by the Route Planner given the new infrastructure. No correction should be made to the calibration functions,

$$P_U^{(1)}(M_1|D), P_U^w(M_2 = A|s_w, D), P_U^b(M_2 = A|s_b, D), \text{ and } P_U^l(M_2 = A|s_l, D).$$

Having no required or desired changes in the calibration functions for a forecast year even though the probabilities, p_R^w, p_R^b, p_R^l , change is an interesting concept. It becomes clearer with some reflection on the application of the calibrated functions to a forecast population. Mode choices are made for each individual/trip in the synthetic population. Therefore, the aggregated D in calibration functions, $P_U^{(1)}(M_1|D)$, $P_U^w(M_2 = A|s_w, D)$, $P_U^b(M_2 = A|s_b, D)$, and $P_U^l(M_2 = A|s_l, D)$ are replaced with I to indicate that the characteristics of the individual such as age, income, and distance to transit are considered. The application of the calibration functions given in Section 2.2 is straightforward.

Consider person I (and all of his individual demographic and transportation characteristics) on a trip or tour. Following the scheme of Section 2.2, mode choice is a two-stage procedure. First, the mode for the trip is determined by a random draw according to the probability function $P_U^{(1)}(M_1|I)$ for $M_1=BI, M, S$ or X . If M_I is chosen to be something other than X , that mode is placed in the activity list. If X is chosen for the M_I mode, the second step is executed. The “mode” X implies that the choice is between auto, walk, bus, and light rail. The mode choice functions, $P_U^w(M_2 = A|s_w, I)$, $P_U^b(M_2 = A|s_b, I)$, and $P_U^l(M_2 = A|s_l, I)$, make the “choice” between auto and walk, auto and bus, and auto and light rail given the characteristics of individual I . The particular choice function used at this stage is picked by the Route Planner. The trip is routed with the general transit mode (T). The Route Planner, using a concept of “rational choice”, determines the choice function by computing the “least time” mode indicated by s_w, s_b , or s_l . The corresponding choice function is utilized and, by random draw from that probability function, the mode A or T is assigned to the activity list. The activity list will have mode A if auto is picked and mode T if not.

The functions $P_U^w(M_2 = A|s_w, I)$, $P_U^b(M_2 = A|s_b, I)$, and $P_U^l(M_2 = A|s_l, I)$ are calibrated as aggregate functions. All travelers with the same characteristics (I) have the same probability (subject to the Route Planner choice s_w, s_b , or s_l) of being assigned modes A or T . The choice of modes when aggregated produce the correct mode splits.

TRANSIMS, on the other hand, is completely disaggregated starting with the generation of the activities and continuing through the microsimulation where individuals and vehicles they occupy are uniquely identified. In this situation, choice functions cannot be

developed that take into account very specific information about each individual traveler (I). For example, given the characteristics (I) and the Route Planner choice of s_w , s_b , or s_l , a bus mode may be the most probable for the trip. However, examination of this trip may show that it is unrealistic. It may, for example, require multiple transfers and long walks between the stops. The analyst may wish to eliminate these types of trips.

It is impossible in the calibration of the choice function to anticipate all of the unusual types of trips that individuals in a population could be assigned. Therefore, rather than try to build checks for all of these in the choice functions $P_U^w(M_2 = A|s_w, I)$,

$P_U^b(M_2 = A|s_b, I)$, and $P_U^l(M_2 = A|s_l, I)$, it is recommended that the mode choice assignment be iterated. Here, unrealistic generalized transit choices denoted by T are replaced with mode A . Trips assigned A are sampled and new modes, either A or T , are determined from the choice functions. This is continued until replacements are obtained for the original mode T that were assigned mode A .

The number of iterations necessary may depend on the fidelity of the user-defined choice functions $P_U^w(M_2 = A|s_w, I)$, $P_U^b(M_2 = A|s_b, I)$ and $P_U^l(M_2 = A|s_l, I)$. This, in turn, requires a trade off between staff time to develop the choice functions and the computer time necessary to iterate the process. Additionally, the time necessary to compute the choice functions is balanced against the time for iteration. Choice functions usually require computation of costs or utilities for all modes under consideration. In one methodology, the costs for both modes, A and T , are computed in the calibration stage. In the application stage, conditional choice functions are used, where costs associated with only one mode are computed. Here, the conditional probability of choosing a particular mode is computed given the knowledge of the costs for the single mode. This is the lowest fidelity possible, and iteration is used to “correct” these choices.

2.2 Mode Choice Using Logits

Logit models are the standard method for mode choice in transportation studies. This subsection outlines logit methodologies for mode choice in TRANSIMS. The method is based on the mathematics shown in the proceeding section. More logit type methodologies will be developed by others for mode choice in TRANSIMS. However, each of these methods will require calibration of $P_U(M|I)$ taking into account the Route Planner behavior as characterized by $P_R(m|M, I)$. The following subsections give details on calibration and application of logits for use in TRANSIMS mode choice.

One method for using logits is to calibrate $P_U(M|I)$ with the mode types BI , M , S , A , and T and assume that the “rational choice” characteristics of the Route Planner for mode T will maintain calibration of s_w , s_b , or s_l —the mode splits for walk, bus, and light rail. It is not clear that the calibration of the mode T can be accomplished with enough fidelity to assure a correct split of s_w , s_b , and s_l . A calibration of this type would simplify the mode split estimation process. The iteration process, which will always be necessary in a completely disaggregated system, may keep these modes in calibration. It would be worthy of future investigations.

In what follows, a two-stage calibration scheme following the procedures in the proceeding subsections is given. There, $P_U(M|I)$ is calibrated for choice among the modes BI , M , S , and X . The second stage is to calibrate the three choice functions represented by $P_R(m|M = X, I)$.

2.2.1 Mode Choice Calibration Using Logits

For the method considered here, walks may be separated into nonessential walks, such as jogs in the morning or walks after dinner. Let these be called w_l . These are distinguished from essential walks, such as walks to work or lunch. In general, these are trips that would have been made using an auto or by transit if they were long enough or if an auto or transit service is available. These are denoted w_2 . Also define b_2 to be a trip containing at least one leg on a bus, and l_2 as a trip with at least one leg on light rail and no legs of the trip on a bus. For example, a sequence of legs on a trip, walk-bus-walk-light rail-walk-bus-walk, is classified as b_2 . Note that b_2 is different that the mode b where a light rail leg would not be allowed and a bus leg is required. The mode l_2 differs from l in that l_2 requires a light rail leg on the trip. Throughout this subsection w , l , and b have the definitions associated with w_2 , l_2 , and b_2 . The Route Planner determines the mode splits for w , l , and b given the activity mode T .

The method defined here requires four logit calibrations in two stages. In the first calibration, the parameters of the calibration function $P_U^{(1)}(M_1|D)$ are determined. This function calibrates the mode splits for the activity modes BI , M , S , W_l , and X . Here, $P_U^{(1)}(M_1|D)$ is taken as a logit function. Equations 2 to 4 show that a mode choice of BI , M , S , or W_l in the activity set will not be changed by the Route Planner. Therefore, this is a standard logit calibration, and the behavior of the Route Planner is not involved.

The second stage of the procedure is to calibrate logit functions for determining the mode split between walk, auto, bus, or light rail given that the first choice model picks the “mode” X . That is, the parameters of the mode choice functions $P_U^w(M_2 = A|s_w, D)$, $P_U^b(M_2 = A|s_b, D)$, and $P_U^l(M_2 = A|s_l, D)$ are estimated to maintain the correct mode split between auto, walk, bus, and light rail. Each of the functions, $P_U^{(1)}(M_1|D)$, $P_U^w(M_2 = A|s_w, D)$, $P_U^b(M_2 = A|s_b, D)$, and $P_U^l(M_2 = A|s_l, D)$ is a logit.

The three second-stage logit fits take into account the “rational choice” behavior of the Route Planner. $P_U^w(M_2 = A|s_w, D)$ represents the choice between auto and walk, given that the Route Planner picks walk when the activity mode is T . $P_U^b(M_2 = A|s_b, D)$ is the logit for the choice between auto and bus, given that the Route Planner picks bus when the activity mode is T . The logit for the choice between auto and light rail, given that the Route Planner picks light rail when the activity mode is T , is $P_U^l(M_2 = A|s_l, D)$.

The procedure assumes that the travel survey is used for the calibration of the logit functions $P_U^{(1)}(M_1|D)$ and $P_U^w(M_2 = A|s_w, D)$, $P_U^b(M_2 = A|s_b, D)$, and $P_U^l(M_2 = A|s_l, D)$. Calibrating the function $P_U^{(1)}(M_1|D)$ is straight forward and requires only that all

essential walks, drives, and transit trips be calibrated as a single mode X . Usual methods for this task using survey data may be employed.

The calibration of the three second-stage functions requires a slightly different techniques. Two possibilities exist. One would use the survey as it is and travel time generated from the survey and external sources as they are in current practice. Separate calibrations would be made for each of the three choices. The second method, and the one recommended, is to locate the survey activities on the base case network. All travel times are then computed by the TRANSIMS Route Planner.

In the second method, driving, bus, essential walking, and light rail trips are isolated. With the survey activities located on the base year network, each X type trip is routed twice—one route is by auto, and the other is with the mode set to T . Each trip is classified as belonging to one of the three groups— s_w , s_b , or s_l —depending on the mode chosen by the Route Planner. Each function, $P_U^w(M_2|s_w, D)$, $P_U^b(M_2|s_b, D)$, and $P_U^l(M_2|s_l, D)$, is calibrated separately by routine logit analysis, where the choice is between auto (A) and the general transit mode (T).

There will be a few trips in the survey where the transit mode (including walks) is changed by the Route Planner. These changes are not expected to be extensive because the Route Planner choice of modes (walk, light rail, and bus) is based on best times.

From the above discussions, it is not clear that logit calibrations of the individual probability functions $P_U^w(M_2|s_w, D)$, $P_U^b(M_2|s_b, D)$, and $P_U^l(M_2|s_l, D)$ keep the aggregated system in calibration. However, Equations 14, 15, and 16 are satisfied with these calibrations. Let a_w , a_b , and a_l be the number of auto trips from the calibration set corresponding to the three sets, s_w , s_b , and s_l . Since the functions are in calibration

$$\begin{aligned} P_U^w(M_2 = A|s_w, D) &= 1 - \frac{w_c}{a_w + w_c} \\ P_U^b(M_2 = A|s_b, D) &= 1 - \frac{b_c}{a_b + b_c} \\ P_U^l(M_2 = A|s_l, D) &= 1 - \frac{l_c}{a_l + l_c} \end{aligned} \quad (18)$$

But, since $a_w + w_c$ estimates the size of s_w in the calibration set,

$$\frac{a_w + w_c}{a_c + w_c + b_c + l_c} = p_R^w$$

or

$$a_w + w_c = p_R^w (a_c + w_c + b_c + l_c) \quad (19)$$

Similar relationships hold for bus and light rail. Substitution of Equation 19 and the similar equations for bus and light rail into Equation 18 yields Equations 14, 15, and 16.

Hence, probability functions $P_U^w(M_2|s_w, D)$, $P_U^b(M_2|s_b, D)$, and $P_U^l(M_2|s_l, D)$ individually calibrated with a logit retain aggregate calibration.

2.2.2 Mode Choice Application Using Logits

Application of logit calibrated functions given in Section 2.2.1 to a forecast year follows the general methodology given in Section 2.1.3. For an individual/trip (I), the choice is made between the “fixed” modes, such as school bus, and the mode X using the choice function $P_U^{(1)}(M_1|I)$. Given the choice (X), the trip is routed with modes A and T . Routing with A ascertains the travel time by auto. Routing with T gives the travel time for the transit mode (which could be walk only) and determines which second stage choice function— $P_U^w(M_2|s_w, I)$, $P_U^b(M_2|s_b, I)$, or $P_U^l(M_2|s_l, I)$ —is used. After this is determined, the mode M_2 is obtained in the usual stochastic manner using the proper second stage choice function.

As noted in Section 2.1.3, the probabilities (or splits) (p_R^w, p_R^b, p_R^l) for the forecast year will most likely not be the same as those in the base year. These changed splits reflect changes in the level of transit service in the forecast year. The change in the level of service determines the non-auto modes on a sequence of legs of a trip, such as walk-light rail-walk-bus-walk. This sequence establishes the correct second step choice function. Travel characteristics that increase transit ridership, such as higher parking and lower transit costs, would be variables in the individual choice functions.

Even with the most carefully calibrated logit choice functions, iteration will most likely be necessary. It was pointed out in Section 2.1.3 that all choice functions are calibrated in an aggregated way. TRANSIMS, on the other hand, eventually follows each individual on a second-by-second basis in the microsimulation. So, an aggregate choice, when applied to a particular individual, may be unrealistic. Such travelers/trips are given the other mode. A replacement traveler/trip is obtained by randomly sampling other travelers/trips until a realistic replacement is found. Replacing trips in the forecast study after they are assigned by the choice function, maintains a calibrated mode split.

2.3 A Second Mode Choice Method

A second methodology does not involve using fitted logit functions. It is a simple methodology based on nonparametric functions of costs that are calibrated to the survey mode splits. Mathematically, it follows the results of Section 2.1, and in that sense is a valid methodology. The methodology presented here is only one of many possible. It is given because of its statistical simplicity, statistical validity, and as a pedagogical tool for those interested in developing different mode choice methodologies. Here, simplified choice functions are used for $P_U^{(1)}(M_1|D)$, $P_U^w(M_2 = A|s_w, D)$, $P_U^b(M_2 = A|s_b, D)$, and $P_U^l(M_2 = A|s_l, D)$.

2.3.1 A Second Mode Choice Method: Preliminary Concepts

Section 2.2 illustrated a methodology where the choice functions $P_U^{(1)}(M_1|D)$, $P_U^w(M_2 = A|s_w, D)$, $P_U^b(M_2 = A|s_b, D)$, and $P_U^l(M_2 = A|s_l, D)$ are calibrated using logit functions. The parameters of these logits are estimated by assigning the activities from the survey to locations on the base case network and determining travel times and the Route Planner choice of mode given that the activity mode choice is T . After accounting for the behavior of the Route Planner, standard estimation procedures are used for the parameters of the logits. Quality of the logit fit is judged statistically and in terms of the traffic and transit trips produced for the base year population.

The method presented here is extremely simplistic, but it demonstrates handling the behavior of the Route Planner along with manipulation of activity files to produce the correct mode splits. In this method, mode splits are calibrated by sampling the base year synthetic population and developing parameterized calibration functions that split the modes. The parameters are chosen so that the sample mode split matches that of target values for the base year. Generally, these target values are derived from the survey.

The calibration functions used are one-parameter functions that compare functions of costs, including travel times, of the various modes. One calibration parameter is determined for each of the functions— $P_U^w(M_2 = A|s_w, D)$, $P_U^b(M_2 = A|s_b, D)$, and $P_U^l(M_2 = A|s_l, D)$. There is a different calibration function for each different set of aggregate characteristics, denoted here by D . Each of these is calibrated to survey counts (a_c , w_c , b_c , and l_c) for the collection of travelers/trips denoted by D .

As with logits, the choice function— $P_U^w(M_2 = A|s_w, D)$ —and the others for bus and light rail, are a function of the costs or utilities for both modes. As an experiment in the application, conditional choice functions were used. These require the costs for only one mode. Estimates of the cost of travel by the other mode and the mode choice itself are determined probabilistically from the single travel cost.

2.3.2 A Second Mode Choice Method: Calibration

A two-stage calibration is assumed in this mode choice methodology. First, the user-defined function $P_U^{(1)}(M_1|D)$ is calibrated to choose among modes $M_1 = BI, M, S$, or X . The mode X is the combined mode A and T . The second stage is the calibration of the functions $P_U^w(M_2 = A|s_w, D)$, $P_U^b(M_2 = A|s_b, D)$, and $P_U^l(M_2 = A|s_l, D)$ given that X is chosen on the first stage.

Each person in the synthetic population for the base year is assigned an activity list. This activity list is a duplicate of the activity list of one of the persons in the activity survey. The TRANSIMS Activity Generator assigns activities and modes by matching household characteristics of the synthetic population and the survey. The locations of the activities for the synthetic individual are picked with a choice function and are not the same as those in the activity survey. However, the modes are not changed by the Activity Generator. With this matching scheme, some the assigned modes (such as bikes, inter-

household shared rides, and school buses) may be reasonable. So, for the methodology presented here, it is assumed that $P_U^{(1)}(M_i|I)$ assigns the modes BI , M , and S to match those given by the Activity Generator. The remainder of the modes (X), which represent auto, walks, and transit, are calibrated separately in the second stage.

Consider only those individuals/trips that are among the walk, drive, or transit trips. The choices for these individuals/trips are between either being assigned the auto mode (A) or the general transit mode (T). This choice is calibrated as follows.

Four cost functions (c_a , c_w , c_b , and c_l) are defined. These denote the “costs” of driving, walking, taking at least one bus leg and possibly light rail legs, and light rail only legs for the trip on the tour. These “costs” may be a function of any variable. This includes travel times, parking costs, transit fares, etc. The costs may also include any information about the traveler such as income and age. Additionally, each cost function must contain at least one tunable or calibration parameter(s) – α . For example, a simple cost function could be

$$Cost = \alpha \log(Income)Time + DollarCost$$

Where, $Time$ is the travel time and $DollarCost$ is the monetary cost of making the trip.

A sample of all of the trips in the “ X ” category are routed twice, once using the auto mode (A) and once using the general transit mode (T). The information necessary to compute the “costs” for each trip is collected. The cost (c_a) is computed from the data, travel times parking costs, household income, etc., determined from the TRANSIMS synthetic population, the TRANSIMS routes, and the TRANSIMS network. Matching the “cost” for auto is one of the costs (c_w , c_b , or c_l) depending on the “choice” of the Route Planner. Information is collected to allow computation of that cost also. The “costs” themselves are computed for various values of the calibration parameter(s) – α . Different values of α give different relative costs between driving and general transit (including walks). The mode choice in the calibration stage is to pick the minimum “cost” mode. The values of α are determined by searching or optimization until the proportion of transit “costs” that are less than the auto “costs” matches the calibration numbers for transit mode splits.

Fig. 2 illustrates this procedure. The relationship between the transit and auto “costs” changes as the calibration parameters change. In Fig. 2, the auto and transit “costs” for two different sets of calibration parameters (α) are shown in the (a) and (b) panels. The arrow in panel (a) shows the general direction in which the points moved as the values of α are changed to produce the points in panel (b). The line in each plot is the equal “cost” line. If the point is above the line, auto has a smaller “cost” than transit. Points below the line favor transit. Values of α are determined so that the proportion of “costs” below the line match the mode split for transit.

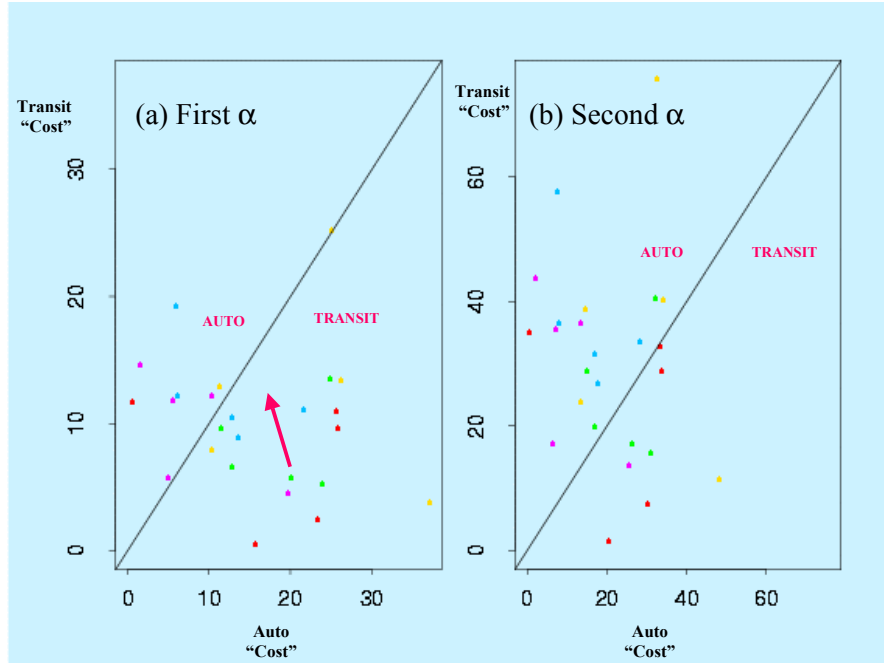


Fig. 2. The relationship between the transit and auto “costs” for two different sets of calibration parameters (α). Values of α are determined so that the proportion of “costs” below the line match the mode split for transit. The arrow shows the general direction the points moved as the values of α changed.

Values of α are determined to match the transit mode split for each the user-defined calibration functions— $P_U^w(M_2 = A|s_w, D)$, $P_U^b(M_2 = A|s_b, D)$, and $P_U^l(M_2 = A|s_l, D)$. The mode split for each of the categories (walk, bus, and light rail) is defined using Equations 14 to 16. For example, consider calibration of the set s_w the walks, (walks are part of the general transit mode T), compared to auto trips. The proportion of transit trips, T , in the category s_w of trips is obtained from Equation 14, as $\frac{w_c}{(a_c + w_c + b_c + l_c)p_R^w}$,

where w_c , a_c , b_c , and l_c are the calibration targets given from the survey or other outside sources. The value p_R^w , as well as p_R^b and p_R^l , is estimated from the results of routing. To estimate these values, a random sample is drawn from those trips classified by $P_U^{(1)}(M_1|I)$ as mode X . These trips are routed with the TRANSIMS Route Planner with mode T . The proportions of these trips that are walks, buses, or light rails are used as estimates of p_R^w , p_R^b , and p_R^l .

This calibration phase determines the exact form of the choice functions— $P_U^w(M_2|s_w, I)$, $P_U^b(M_2|s_b, I)$, or $P_U^l(M_2|s_l, I)$ —that are applied to the individuals and trips in the forecast year.

2.3.3 A Second Mode Choice Method: Application to a Forecast Year

For application, the choice functions $P_U^{(1)}(M_1|I)$, $P_U^w(M_2|s_w, I)$, $P_U^b(M_2|s_b, I)$, and $P_U^l(M_2|s_l, I)$ have been calibrated to the base year. That is, the calibration parameters (α) have been determined. Hence, the “costs” for travel by either mode A or T can be calculated. Three procedures for determining the assigned modes are discussed here. Each of these methods requires that a sample of trips/tours from the forecast year be routed by both modes, auto (A) and general transit (T). The assignment of modes by the first of these methods requires that all trips/tours be routed by both auto (A) and general transit (T). The second two statistically valid procedures for assigning modes to the forecast activity set require routing by general transit (T) only or by auto (A) only.

The mode at the first stage of the assignment is given from the function $P_U^{(1)}(M_1|I)$. If the chosen mode at this stage is X , then the second stage choice is between modes A and T . To determine the mode in this second stage, a sample of the trips/tours is routed by both general transit (T) and auto (A). The “costs” for transit (c_t) and for auto (c_a) are computed and plotted as shown in Fig. 3. The following three methodologies for assigning modes to the forecast year all require this type of sample.

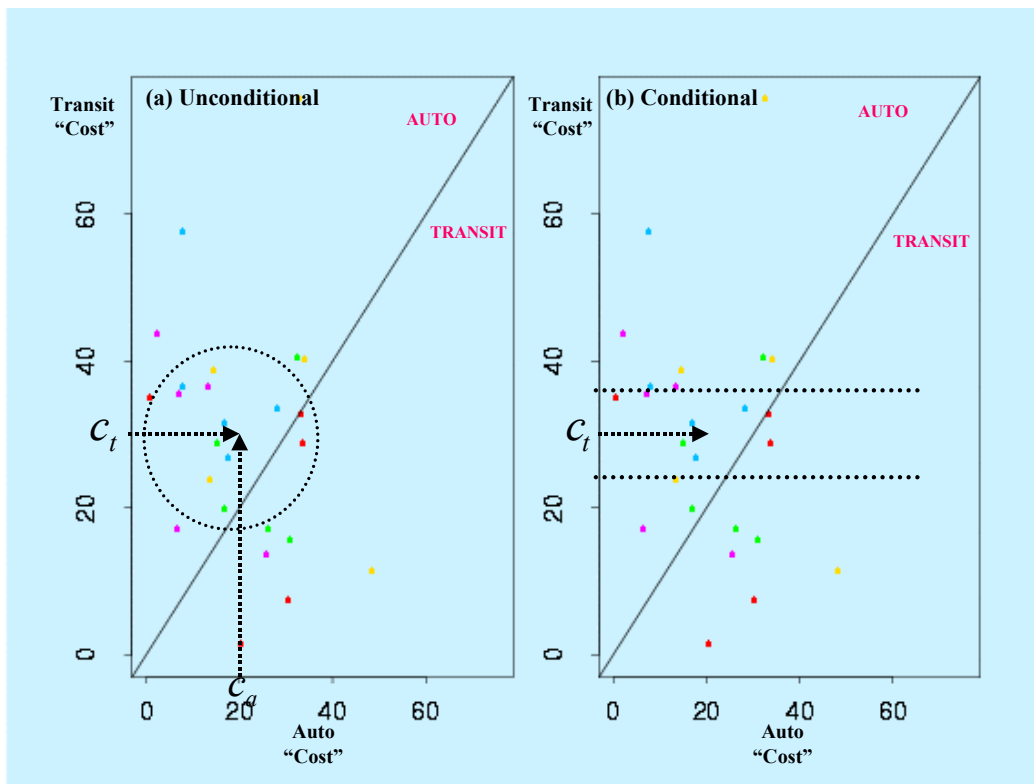


Fig. 3. Two techniques for stochastic mode selection. Panel (a) shows the method when both “costs” are computed. Estimates of mode selection using panel (b) require only generalized transit “costs.”

2.3.3.1 Application Routing by Transit and Auto

The first methodology requires that all trips/tours be routed by both general transit (T) and auto (A). There are two approaches for deciding between A and T given the two costs, c_a and c_t . The first is a simple comparison of the two. If $c_a < c_t$, then A is assigned to the activity list. Otherwise, T is assigned. This method is completely deterministic. The implicit assumptions are that the “cost” functions reflect the thought processes of the traveler and that the traveler will always pick the lowest cost alternative when choosing between auto and general transit. Both of these assumptions are most likely not true.

The calibration of the second stage choice functions is an aggregate procedure. No aggregate procedures, including logits, make the assumption that the choice functions are precise enough to make the choice for A based on the relationship $c_a < c_t$ alone. The relationship $c_a < c_t$ is only an indication that mode A is more likely. Therefore, almost all choice functions are a stochastic function of the “costs” c_t and c_a .

A stochastic approach is possible given the two costs— c_t and c_a —and a sample from the forecast year of trips routed on both modes. This method is shown in Fig. 3(a). The data points on the plot represent the calibration “cost” data computed with the estimated parameters (α). The point (c_a, c_t) is located on the plot. This is shown in Fig. 3(a) as the intersection of the two dashed arrows. A circle of radius r is drawn with the point (c_a, c_t) as the center. The proportion of points in the circle above the equal cost line becomes the probability of choosing mode A .

In this second stochastic approach, the radius of the circle should be large to make some reasonable estimates of the probabilities. Each estimated probability should probably be based on at least 10 points. The radius r does not need to be constant for all choices. Given a point (c_a, c_t) , the minimum r containing 10 points could be used. The probabilities based on these computations may be computed and tabled before the mode choice exercise. Any regular shape around the point (c_a, c_t) could be used to compute these probabilities. A rectangle could make these computations easier.

It is easily shown that as the area selected around the point becomes extremely small or extremely large, the procedure remains in calibration. Given the stochastic nature of the calibration, a small area around the chosen point will have little effect.

2.3.3.2 Application Routing by Transit Only

The second methodology is conditional on only the transit cost (c_t). After routing with mode T , the cost (c_t) and the appropriate choice function are determined. At this point, the sample forecast data with the time “costs” for both modes is plotted as before. An area of the plot is captured where the transit “costs” are between $c_t \pm \Delta c_t$. This is shown in Fig. 3(b). The computation of choice probabilities is the same as the previous method. The proportion of calibration points between $c_t \pm \Delta c_t$ above the equal cost line is the estimated probability of driving.

The above method maintains calibration. The proportion of points above the line in the area $c_t \pm \Delta c_t$ estimates the conditional probability $P(C_a < C_t | C_t = c_t)$. If $f(c_t)$ is the probability function representing the choice of c_t , then

$$P(C_a < C_t) = \int P(C_a < C_t | C_t = c_t) f(c_t) \partial c_t$$

2.3.3.3 Application Routing by Auto Only

The third methodology where only the auto “cost” (c_a) is computed is more complex. When a trip is routed with mode T , one of the three second-stage choice functions is automatically picked. After routing by mode A , three auto “costs” (c_a^w, c_a^b, c_a^l)—one for walks, one for buses, and one for light rail as alternatives—can be computed. The actual “cost” and choice function to use is unknown. However, a statistically valid method using these computed costs is developed below.

In the forecast year, a random sample of X mode trips has been routed by transit (T). These samples are used to form the scatter plots as shown in Fig. 3(b), and for each of the three choice functions, a conditional probability of choosing auto given c_a is computed using a similar methodology to that given above. Let these probabilities be denoted by $P^w(M_2 = A | c_a^w, s_w)$, $P^b(M_2 = A | c_a^b, s_b)$, and $P^l(M_2 = A | c_a^l, s_l)$, where (c_a^w, c_a^b, c_a^l) are the computed cost functions for the three calibration functions, and (s_w, s_b, s_l) denotes that the Route Planner chooses the corresponding calibration function. Then, the probability that auto (A) is chosen for the mode given $C = (c_a^w, c_a^b, c_a^l)$ is

$$\begin{aligned} P(M_2 = A | C) &= P^w(M_2 = A | c_a^w, s_w) P(s_w | C) + \\ &P^b(M_2 = A | c_a^b, s_b) P(s_b | C) + \\ &P^l(M_2 = A | c_a^l, s_l) P(s_l | C) \end{aligned}$$

Here $P(s_w | C)$, $P(s_b | C)$, and $P(s_l | C)$ are the probability the Route Planner would pick walk, bus, or light rail for a particular trip/tour given only information on the auto “costs”, $C = (c_a^w, c_a^b, c_a^l)$.

The probabilities $P(s_w | C)$, $P(s_b | C)$, and $P(s_l | C)$ are estimated using the sample data from the forecast year. If the analyst believes that knowing the auto “cost” functions for a particular trip/tour supplies no information about the behavior of the Route Planner, then the proportion of trips assigned to walk, bus, and light rail ($p_{R_f}^w, p_{R_f}^b, p_{R_f}^l$) may be used as estimates of these probabilities. In many studies however, this will not be true. If the “cost” functions are a function of the distance traveled, then one would suspect that the Route Planner would be more likely to “choose” walk for the shorter trips and transit for the longer trips. In this case, the Route Planner “choice” of walk or transit is not independent of the particular costs (C).

There are numerous techniques to estimate the probabilities $P(s_w | C)$, $P(s_b | C)$, and $P(s_l | C)$. Three are given here. To simplify the discussion, we consider only two partitions of the routes—those that contain only walk legs and those that contain at least one transit leg. Let the corresponding probabilities and costs be $P(s_w | C)$, $P(s_t | C)$, and $C = (c_a^w, c_a^t)$.

The first method for estimating $P(s_w | C)$ and $P(s_t | C)$ is shown in Fig. 4. The black points in the figure are the calibration data where the Route Planner produces all walk legs. The red points represent those where the route from the Route Planner has at least one transit leg. The arrows represent the auto costs, $C = (c_a^t, c_a^w)$, for a trip/tour in the forecast data set. These costs are computed using the calibration parameters determined in the calibration step. The probabilities $P(s_w | C)$ and $P(s_t | C)$ are estimated by determining the proportion of black or red dots in a circle centered at $C = (c_a^t, c_a^w)$.

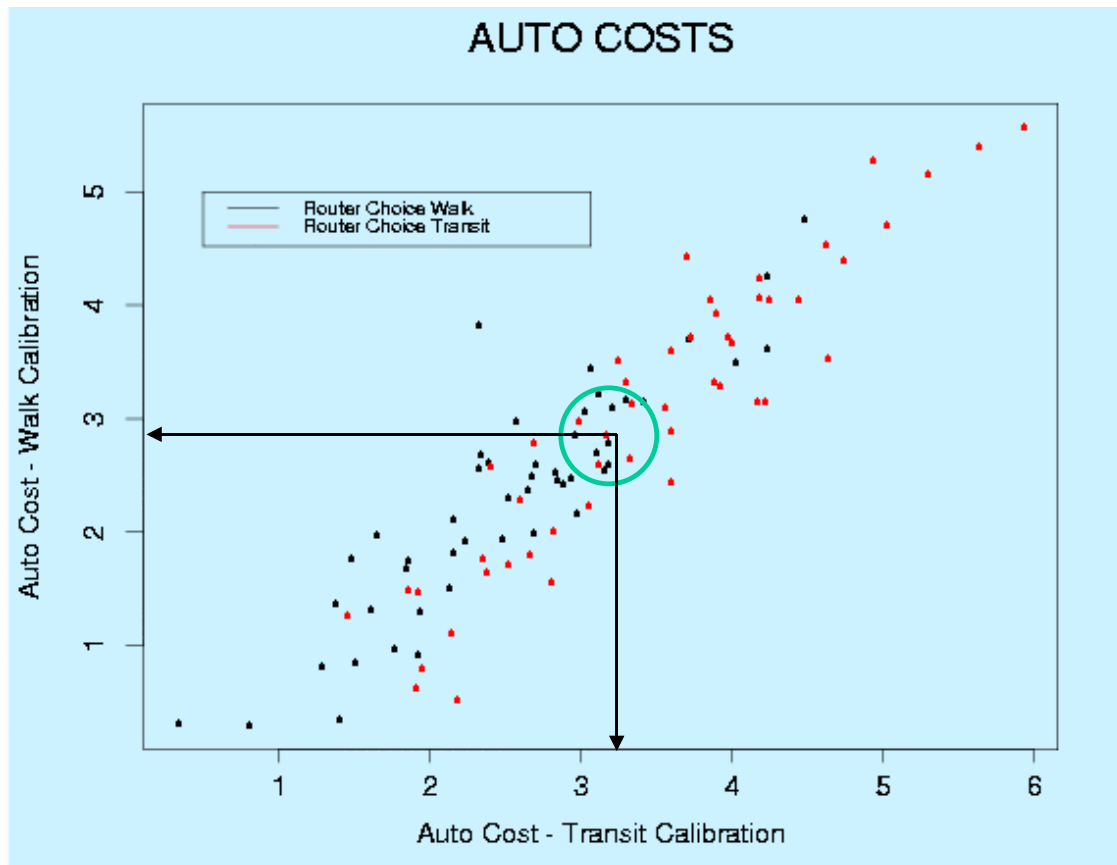


Fig. 4. Auto costs for transit and walk calibrations plotted against each other. The black points are the calibration data where the Route Planner produces all walk legs. The red points represent those where the route from the Route Planner has at least one transit leg. In an application, the auto costs $C = (c_a^t, c_a^w)$ are computed, and $P(s_w | C)$ is estimated by the proportion of black dots in the circle.

If the above method is used to estimate the probabilities, the sample size for the application set must increase as the number of possible mode choices, (e.g., s_w, s_t) increases. For example, in the case of three possibilities, the circle in Fig. 4 becomes a sphere and more points would need to be captured to make reasonable estimates of the probabilities.

A second technique for estimating the probabilities is to fit a simple logit type model to the data in Fig. 4. That is, a choice model is fit as a function of the two costs, and the probabilities given by

$$P(s_w | C = (c_a^t, c_a^w)) \propto \exp(\beta_0 + \beta_1 c_a^t + \beta_2 c_a^w)$$

where $(\beta_0, \beta_1, \beta_2)$ are estimated parameters.

A third possibility is to identify surrogate variable(s) (V) for the costs, here $C = (c_a^t, c_a^w)$, and use these variable(s) to determine the relative likelihood of the choices (e.g., s_w, s_t). It is assumed that the dimension of V is less than that of C . For example, a variable such as distance may be a substitute for the two costs. If true, then the probabilities $P(s_w | V)$ and $P(s_t | V)$, where V is the substituted variable, stand in place of the probabilities $P(s_w | C)$ and $P(s_t | C)$. Logits may be fit to the variable(s) V . However, V is composed of a single variable, then estimates of the probabilities may be obtained by moving windows as in Fig. 4, but in this case in one dimension.

2.3.3.4 Application Summary

The methods given above resemble the logit methodologies in the previous subsections. They are all aggregate methods that are applied to individuals. The same iterations outlined in the logit sections are also necessary here.

The conditional methodologies where the choice is based on either the transit “cost” alone or the auto “cost” alone are interesting. The transit “cost” methodology has less fidelity than knowing both “costs”. Knowing the auto “cost” alone has less fidelity than knowing the transit “cost” alone. Each of the methods are statistically valid in that they maintain calibration, but the decreasing fidelity may require more iterations. In TRANSIMS, routing by transit (T) is slower than routing by auto (A). The trade off between iterations and fidelity is interesting and will undoubtedly be studied in the future.

The conditional methodologies discussed above can also be applied to logits if the logit functions can be written as a separable combination of two functions, one for mode T and one for mode A . In this case, a sample of the forecast data can be arranged as shown in Fig. 2 and Fig. 3. Given a logit type “cost” for one of the modes, the surrounding points on the plot can be captured. The probability of choosing auto (or transit) computed from the logit function is assigned to each point. The probability of choosing auto is then the sum of the probabilities in the auto region divided by the sum of all the probabilities captured.

3. SUMMARY

Mode calibration in TRANSIMS must take into account the “rational choice” behavior of the Route Planner. Modes are assigned in the activity list for each trip. If this assignment is made with a mode choice algorithm that was calibrated without taking into account the behavior of the Route Planner, the resulting mode splits in the Traffic Microsimulator will be incorrect. In this case, the number of walk trips will be overrepresented, while the numbers of transit and drive trips will be underrepresented.

A methodology for calibrating mode choice taking into account the behavior of the Route Planner is given in this document. This methodology makes mode choices in two steps. The first step, is a calibration between those modes that will not be changed by the Route Planner and all those modes that potentially could change (walk, drive, and transit). The second stage calibration determines the mode choice for those modes that the Route Planner could change. It should be noted that this methodology is not the only method for calibrating mode choice in TRANSIMS.

In most mode choice applications, travel times and costs are determined for each trip and each mode of transportation. Some methodologies are given in this document that allow for statistically correct procedures with routing by only one mode. These methodologies are independent of the form of the calibration function.

The method of calibration is the analyst’s choice. Full logit functions that allow for choice between all modes for each traveler and each type of trip may be used. In addition, simpler logit type calibration functions may be calibrated and used to assign modes for subsets of the population. For example, a subset may be those travelers who work downtown and live near a transit stop.

TRANSIMS is a completely disaggregate system. All statistical methods, such as logit fits or the more simple fits described above, are aggregate procedures. When these procedures are applied to a disaggregate population, a portion of the travelers will be assigned nonrealistic modes. For example, an assigned transit trip may require five or six transfers. If there are many of these nonrealistic trips, the mode choice for the trips must be changed by iteration.